



## Short but Weak: The Z-DNA Lone-Pair... $\pi$ Conundrum Challenges Standard Carbon Van der Waals Radii

Holger Kruse, Klaudia Mrazikova, Luigi d'Ascenzo, Jiri Sponer, Pascal Auffinger

### ► To cite this version:

Holger Kruse, Klaudia Mrazikova, Luigi d'Ascenzo, Jiri Sponer, Pascal Auffinger. Short but Weak: The Z-DNA Lone-Pair... $\pi$  Conundrum Challenges Standard Carbon Van der Waals Radii. *Angewandte Chemie International Edition*, 2020, 59 (38), pp.16553-16560. <10.1002/anie.202004201>. <hal-02953786>

**HAL Id: hal-02953786**

**<https://hal.science/hal-02953786v1>**

Submitted on 6 Oct 2020

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



HAL Authorization

# Short but weak! The Z-DNA lone-pair... $\pi$ conundrum challenges standard carbon vdW radii

Holger Kruse,<sup>\*,[a]</sup> Klaudia Mrazikova,<sup>[a,b]</sup> Luigi D'Ascenzo,<sup>[c]</sup> Jiri Sponer,<sup>[a,d]</sup> and Pascal Auffinger,<sup>\*,[e]</sup>

- [a] Dr. Holger Kruse, Klaudia Mrazikova, Dr. Jiri Sponer  
Institute of Biophysics of the Czech Academy of Sciences  
Královopolská 135, 61265 Brno, Czech Republic  
E-mail: kruse@ibp.cz
- [b] Klaudia Mrazikova  
National Centre for Biomolecular Research, Faculty of Science,  
Masaryk University  
Kamenice 5, 625 00 Brno, Czech Republic
- [c] Dr. Luigi D'Ascenzo  
Department of Integrative Structural and Computational Biology,  
The Scripps Research Institute  
La Jolla, CA 92037, USA
- [d] Dr. Jiri Sponer  
Centre of Advanced Technologies and Materials, Faculty of Science,  
Palacký University  
tř 17 Listopadu 12, 771 46, Olomouc, Czech Republic
- [e] Dr. Pascal Auffinger  
Université de Strasbourg, CNRS, Architecture et Réactivité de l'ARN, UPR9002,  
Strasbourg, France  
E-mail: p.auffinger@ibmc-cnrs.unistra.fr

Supporting information and the ORCID identification number(s) for the author(s) of this article is given via a link at the end of the document.

**Abstract:** Current interest for lone-pair... $\pi$  (lp... $\pi$ ) non-covalent interactions is gaining momentum in biochemistry and (supramolecular)chemistry. However, the physicochemical origin of the exceptionally short ( $\approx 2.8$  Å) oxygen-to-nucleobase plane distances observed in prototypical Z-DNA CpG steps remains unclear. High level quantum mechanical calculations with SAPT2+3 interaction energy decompositions demonstrate that lp... $\pi$  contacts do not result from  $n \rightarrow \pi^*$  orbital overlaps but from weak dispersion and electrostatic interactions combined with stereochemical effects imposed by the highly strained structural context. They also suggest that the carbon van der Waals (vdW) radii originally derived for  $sp^3$  carbons (1.70 Å) should not be used for the smaller  $sp^2$  carbons bonded to electron-withdrawing groups. With a more adapted carbon vdW radius, these lp... $\pi$  contacts no longer enter the sub-vdW category. Our findings challenge the whole lp... $\pi$  concept that refers to elusive orbital interactions that fail to explain the experimentally derived short interatomic contact distances.

These distances are among the shortest distances of approach of an oxygen atom to an aromatic ring observed in chemical and biochemical systems. In Z-DNA, lp... $\pi$  contact distances ( $< 3.0$  Å) are shorter than the average stacking distance between nucleobases in B-DNA (3.3–3.4 Å).<sup>[3]</sup> They are also shorter than the often-cited (3.22 Å) sum of the van der Waals (vdW) radii of an oxygen (1.52 Å) and a carbon atom (1.70 Å)<sup>[2,13]</sup> placing them *de-facto* into the sub-vdW category of NCIs. This criterion, i.e. a contact distance, between two atoms, that is shorter than the sum of their vdW radii, is frequently used to assess the attractive character of short interactions.<sup>[5b,6a,6f,7,14]</sup>

## Introduction

Lone-pair... $\pi$  (lp... $\pi$  or  $n \rightarrow \pi^*$ ) interactions typically involve the stacking of an oxygen atom with an aromatic ring. They constitute one of the most recent additions to the non-covalent interaction (NCI) family involving  $\pi$  systems.<sup>[6]</sup> While considered as promising in supramolecular<sup>[5b,6f,7]</sup> and catalytic chemistry,<sup>[8]</sup> their physico-chemical characteristics are still debated. Quantum mechanical (QM) calculations established that these NCIs are weakly attractive.<sup>[6a,6c,6f-i,9]</sup> However, no rational explanations have emerged regarding the exceptionally short oxygen to guanine (G) nucleobase plane contact distances occurring in the prototypical Z-DNA CpG steps<sup>[11]</sup> (also called Z-steps<sup>[5a,12]</sup>, **Figure 1**).

**Figure 1. Lp... $\pi$  model systems.** **a)** Oxygen-to-guanine lp... $\pi$  contact observed in a CpG Z-step extracted from a Z-DNA d(CpGpCpGpCpG)<sub>2</sub> duplex X-ray structure.<sup>[4]</sup> The CpG Z-step is similar to those present in UNCG RNA tetraloops.<sup>[5]</sup> The ribose O4' atoms are yellow; the cytosine (C) and guanine (G) carbon atoms are wheat and white, respectively; the lp... $\pi$  stacking contact is represented by a dashed line (see **Figure S1** for broader Z-DNA and UNCG contexts and **Figure S2** for a DME...G and d-ribose...G system description). The experimental O4' atom to the guanine plane distance is 2.85 Å; the shortest oxygen contact distance to the C2 atom is 2.95 Å (**Table S1**). The average experimental O4'...plane and O4'...C2 distances over all six CpG steps are  $2.83 \pm 0.05$  and  $2.94 \pm 0.05$  Å, respectively. **b)** Depiction of the localized intrinsic bond orbitals (IBO)<sup>[10]</sup> of a lp... $\pi$  model system involving dimethylether (DME) and a G nucleobase. The isosurface encloses 80% of the orbital's density. The calculations used a B3LYP/def2-QZVP wavefunction. **c)** d-ribose...G model.

Among other authors,<sup>[15]</sup> Jack Dunitz and Angelo Gavezzotti<sup>[16]</sup> (D&G) have challenged the use of this criterion. D&G argued that short interatomic distances in crystals and biomolecular systems do not always warrant attractive interactions nor significant orbital overlaps as implied by the "lp... $\pi$ " term. Instead, these authors proposed that short distances in crystals may result from incentives to minimize the free energy of the system by close-packing optimization.<sup>[15a,16a,16c]</sup>

These structural effects, sometimes called 'compression effects',<sup>[15a,17]</sup> are generally associated with weak interactions characterized by relatively shallow energy profiles. In such contexts, proximal forces can push two molecular fragments together at distances closer than the minimum interaction energy distance of the involved pairs of atoms or at distances below the alleged sum of their vdW radii.<sup>[17]</sup> Herein, we address the following: can weakly attractive lp... $\pi$  interactions induce by themselves the short stacking distance observed in CpG Z-steps (< 3.0 Å) or are there idiosyncratic Z-DNA factors at play?

To tackle this issue, we performed calculations at the highest QM level applied to date to a lp... $\pi$  system. As an initial model, we used the dimethylether-guanine complex (DME...G; [Figure 1b](#)). To add more insights to the lp... $\pi$  contact description, we performed also calculations on a deoxyribose...guanine (*d*-ribose...G) system ([Supporting Information; Figure 1c](#)). Finally, to appraise the influence of the molecular environment on the lp... $\pi$  geometry, we optimized a full Z-DNA duplex.

Our findings lead us to speculate that the Z-DNA structural context is operative in pushing the deoxyribose oxygen atom close to the nucleobase plane in proximity to the van der Waals (vdW) depression of the guanine pyrimidine ring. Thereby, we disagree with previous interpretations suggesting that the formation of short lp... $\pi$  contacts is facilitated by a specific polarization of the guanine in a Z-DNA environment involving the presence of magnesium ions.<sup>[6a,6c,18]</sup> In addition, our investigations led us to challenge the validity limits of the 1.70 Å carbon vdW radius<sup>[2,13]</sup>, often unduly used to categorize short lp... $\pi$  contacts into attractive NCIs, especially when  $sp^2$  carbons bound to electron-withdrawing groups are involved as in a guanine nucleobase.

## Results

### Closest oxygen...guanine contact in a DME...G system.

The closest-contact surface —i.e., the surface at which the interaction energy is at its minimum— was calculated by moving DME as a rigid probe over the guanine plane ([Figure 2](#)). For that, we used dispersion-corrected DFT at the highly accurate DSD-BYLP-D3/ma-def2-QZVPP (DHDF-D3) theory level —see [Methods](#).

The closest-contact position of the O4' atom over the guanine plane almost overlaps with the guanine pyrimidine ring centroid ([Figure 2](#)). The 2.97 Å contact distance is close to the average distances of O4' atoms over guanine nucleobases derived from a Protein Data Bank (PDB) survey of nucleic acid crystallographic structures (2.9 Å and 3.0 Å for DNA and RNA, respectively).<sup>[5a]</sup> The calculated interaction energy between the two molecular fragments at this position is  $-3.85 \text{ kcal.mol}^{-1}$

**Figure 2.** DME...G and *d*-ribose...G surfaces. **a)** Closest-contact surfaces. **b)** Interaction energy surfaces. Note that the closest-contact distance is not systematically associated with the strongest interaction energy as the balance of interactions differs depending on the local environment of the probed point. **a, b)** Top, representation of the vertical scan-axis (dashed line) over the pyrimidine centroid for the DME...G and *d*-ribose...G systems. Guanine ring atoms are labelled according to standard conventions. Red dots and dark blue arrows mark interpolated closest-contact and energy minima of the guanine pyrimidine and imidazole rings. Black dots mark basepoints used for the DHDF-D3 calculations —see [Methods](#). The nucleobase is located in the  $z=0$  plane. An isoline stepping of 0.02 Å and  $0.25 \text{ kcal.mol}^{-1}$  was respectively used for the closest-contact and interaction energy surfaces. For the DME and *d*-ribose orientations used for the calculations, see [Figure 1b](#). The yellow star corresponds to the crystallographic average position derived from [Figure 3b](#).

([Figure 2](#)). This interaction energy is small when compared to the already small CpG intra-strand nucleobase-stacking interaction energy ( $\approx -6 \text{ kcal.mol}^{-1}$ ) and that of a G=C base pair ( $\approx -30 \text{ kcal.mol}^{-1}$ ) as derived from QM calculations conducted at the same DHDF-D3 level.<sup>[19]</sup>

A second shallower minimum (3.04 Å) is located over the imidazole ring close to its centroid. However, no oxygen...imidazole contacts are observed in Z-DNA ([Supporting information](#) and [Figure 3](#)).

**Statistical overview of oxygen...guanine stacking contacts in Z-DNA CpG steps.** To get experimental data of higher quality than those mentioned above that used a tolerant crystallographic resolution limit of 3.0 Å,<sup>[5a]</sup> we collected an ensemble of 114 CpG Z-steps extracted from high-resolution ( $\leq 1.0$  Å) Z-DNA structures. We obtained a slightly more precise lp... $\pi$  average distance of  $2.86 \pm 0.06$  Å instead of  $2.9 \pm 0.1$  Å ([Figure 3](#)). In both ensembles, most of the points cluster in between the center of the pyrimidine ring and the C2 atom but not above the pyrimidine ring centroid as for DME...G. Interestingly, the experimental O4'...C2 distance ( $2.98 \pm 0.06$  Å; [Table S1](#)) is 0.24 Å and 0.34 Å shorter than the 3.22 Å sum of the oxygen (1.52 Å) and carbon (1.70 Å) standard vdW radii and the equivalent  $d(\text{O4}'\dots\text{C2})=3.32$  Å distance in DME...G, respectively. These values underscore clear geometrical differences between the experimental and calculated systems. The following aims at understanding these differences and the nature of the lp... $\pi$  interaction.

**Figure 3.** Comparison between two ensembles of O4' to guanine stacking contacts extracted from the PDB. a) The first ensemble, derived from a previous study on  $lp\cdots\pi$  contacts involving DNA X-ray structures (PDB; February 2016) with resolutions  $\leq 3.0$  Å,<sup>[5a]</sup> comprises 193 O4' atoms. b) The second ensemble, derived from Z-DNA X-ray structures, comprises 114 O4' atoms (PDB; November 2019; hydrogens not shown; resolutions  $\leq 1.0$  Å). Both figures provide top views (5'-side) and their 90° rotation.

**Nature of the  $lp\cdots\pi$  contact in the DME...G system.** To investigate the bonding nature of the  $lp\cdots\pi$  contact, we have carried out SAPT calculations at a very high level of theory (see [Methods](#)).<sup>[20]</sup> SAPT decomposes the interaction energy into dispersion, electrostatics, exchange-repulsion and induction contributions.<sup>[20]</sup> SAPT calculations were performed for each point of the vertical DME...'guanine pyrimidine ring centroid' scan ([Figure 4](#)). We found that the dispersion energy ( $\Delta E_{\text{disp}}$ ) is the dominant attractive energy component. At the minimum (2.97 Å) of the 'total energy',  $\Delta E_{\text{disp}}$  is roughly compensated by the exchange-repulsion term ( $\Delta E_{\text{exch}}$ ) that augments below 3.5 Å as a result of overlapping fragment-charge-densities. In the same distance range, the electrostatic contribution ( $\Delta E_{\text{elect}}$ ) becomes more attractive due to stabilizing 'charge-penetration' effects.<sup>[21]</sup> Negligible contributions near the energy minimum are provided by the induction ( $\Delta E_{\text{ind}}$ ) and charge-transfer ( $\Delta E_{\text{CT}}$ ) terms ([Table S2](#)).

A partial covalent character associated with an  $n\rightarrow\pi^*$  stabilization would have resulted in both, a much larger charge-transfer and induction contributions, the latter being caused by the wavefunction interference (also called 'resonance' in the valence bond language).<sup>[22]</sup> Thus, the DME...G energy decomposition profile is typical for systems involving NCIs without notable orbital interactions such as nucleic acid base stacks or benzene dimers.<sup>[23]</sup> At a different level of theory, natural bond orbital (NBO)<sup>[24]</sup> computations point to 'non-significant' orbital interactions ( $lp\cdots\pi$  NBO interactions in DME...G is  $-0.13$  kcal.mol<sup>-1</sup>; for a discussion on "non-significant" NBO values, see [Supporting Information](#)).

**A more comprehensive d-ribose...G model shows similar trends.** To address the geometrical differences mentioned above, we performed calculations on a d-ribose...G system ([Figure 1c](#)). This model (2.93 Å) comes closer to the Z-DNA  $lp\cdots\pi$  contact distance (2.86 Å) than the DME...G model (2.97 Å). It indicates that the effects induced by the additional ribose atoms are not to be neglected, but it exaggerates the in-plane shift of O4' towards C2 ([Figure 2](#)).

Contrary to the  $d(O4'\cdots C2)=3.28$  Å in DME...G, that is above the estimated 3.22 Å sum of the carbon and oxygen vdW

radii,<sup>[2,13]</sup>  $d(O4'\cdots C2)=3.01$  Å in d-ribose...G suggests a substantial vdW overlap. However, like for DME...G, the d-ribose...G SAPT analysis does not support meaningful orbital interactions ([Figure 4](#); [Supporting Information](#)). This absence of evidence of any significant orbital interactions in both systems prompted us to examine other factors explaining the short experimentally derived O4'...C2 and O4'...guanine distances.

**Is the standard carbon vdW radius appropriate for describing short  $lp\cdots\pi$  NCIs?** At this point, we questioned the reliability of the isotropic carbon vdW radius value (1.70 Å)<sup>[13]</sup> that is widely used to assign  $lp\cdots\pi$  contacts to the sub-vdW class of interactions. Even though vdW radii are not physical observables, they hold a strong place as a chemical concept. Yet, the fact that electron density clouds associated with atomic nuclei are anisotropic and affected by the chemical environment is often neglected.<sup>[13,25]</sup> To estimate the effective volumes of guanine atoms, we calculated  $\alpha(0)$  atom-in-molecule (AIM) polarizabilities along with partial atomic charges ([Table S3](#); note that atomic polarizabilities are sometimes considered as equivalent to atomic volumes<sup>[26]</sup>). We found that among the guanine ring atoms, the  $\alpha(0)$  value of the C2 atom is smaller than those of the neighboring atoms. Therefore, C2 appears to be the smallest guanine nucleobase atom, a fact that helps to understand the short experimental O4'...C2 contact distance.

**Figure 4.** Interaction energy decomposition (SAPT) analysis for a) DME...G and b) d-ribose...G along the normal to the nucleobase plane. The DME...G curve is plotted over the pyrimidine centroid while the d-ribose...G is plotted over the DHDF-D3 surface minimum. The DHDF-D3 interaction energy curve is shown for comparison (dashed line). Inlays mark the minima derived from the DHDF-D3 surfaces and the corresponding DHDF-D3 interaction energy values. The small difference between the DHDF-D3 ( $-3.85$  kcal.mol<sup>-1</sup>) and the "less accurate" SAPT interaction energy ( $-3.22$  kcal.mol<sup>-1</sup>) at the minima of the total interaction energy curves, is of no consequence for the interpretation of the SAPT energy decomposition ([Table S2](#)).

**Figure 5. Guanine electron density maps and visualization of the pyrimidine ring depression.** a) Slice of the electron density around the G4 residue (PDBid: 3P4J) derived from a transferable aspherical atom model (TAAM) refinement<sup>[1]</sup> at 3.0 (blue), 1.0 (red) and 0.1 (grey) sigma ( $\sigma$ ). Note the red density that reveals the presence of hydrogen atoms and underlines the high quality of the experimental data. b) Slice of the deformation electron density map, at +0.05 (blue) and -0.05 (red) sigma. c) CPK model of a guanine. A blue arrow marks the vdW depression in the middle of the pyrimidine ring. Note that in CPK models, the carbon atom volume is larger than those of nitrogen and oxygen atoms since they are based on conventional vdW radii.<sup>[2]</sup> d) Computed electron density contours through the intersecting C2...C5 plane. Contour values range from 0.01 (blue) to 0.001 (red). An isovalue of 0.001 is typical for representing a vdW-like surface. An enclosing grey background is used for clarity.

When compared to the indole phenyl ring carbon atoms, the C2 atom  $\alpha(0)$  value is shrunk by almost 20%. This effect can be explained by the polarizing (electron-withdrawing) nature of the attached nitrogens displaying negative partial charges (N1/N2/N3  $\approx$  -0.33/-0.50/-0.43) that contrast with the positive partial charge of the C2 atom (+0.39; **Table S3**). We infer that this large  $\approx$ 0.7/0.9 electron charge difference between C2 and the connected nitrogens significantly affects their respective atomic volumes as expressed by the calculated  $\alpha(0)$  values (**Table S1**).

To get further insights into the Z-step guanine electronic

structure, we visualized data from a recent transferable aspherical atomic model (TAAM) refinement performed on the ultrahigh resolution (0.55 Å) 3P4J Z-DNA crystal structure. This refinement takes into account the loss of atom neutrality as well as aspherical density functions.<sup>[1]</sup> **Figure 5** shows slices of the electron density and deformation maps around the guanine nucleobase. These slices highlight significant differences between the C2 atom and the connected N1/N2/N3 nitrogens that are related to the calculated partial charges mentioned above (**Table S3**). Although these maps cannot provide an estimate of the vdW radii of the concerned atoms, they do not support a carbon vdW radius much larger than those of the neighboring nitrogens. At this stage, it appears that a 1.70 Å vdW radius for the guanine  $sp^2$  carbon (C2) is a severe overestimation of its effective radius, therefore challenging the sub-vdW status of the  $lp \dots \pi$  interaction in Z-DNA.

**Is the pyrimidine ring vdW depression involved in the short O4'...nucleobase distance?** Next, we need to rationalize the short experimental O4'...plane distance ( $2.83 \pm 0.05$  Å; **Figure 1**). For this, we noted that CPK models of purine and pyrimidine rings feature, like any other aromatic ring, a rarely addressed small but distinctive vdW depression in their center<sup>[27]</sup> that is also observable in the electron density maps (**Figure 5a**). Such a depression allows the fit of an oxygen atom (**Figure 5c**) as nicely emphasized by a two-dimensional cut of a DFT density (B3LYP/def2-QZVP) through the DME...G pyrimidine ring centroid (**Figure 5d**). From these representations, it is understandable, although difficult to quantify, that the O4' atom involved in the  $lp \dots \pi$  contact benefits from this depression to achieve a short O4'...plane distance.

**Short Z-DNA  $lp \dots \pi$  contacts involve stereochemically induced “compression effects”.** Finally, as we established that short contact distances are not due to specific orbital interactions and that associated NCIs are not strong enough to push the O4' atom towards the nucleobase, it remains to understand why such short contacts are observed in Z-DNA. For that, we performed a full geometry optimization of the  $d(CpGpCpGpCpG)_2$  Z-DNA crystal structure (**Figure S1**),<sup>[4]</sup> by using dispersion-corrected DFT (DFT-D3 = B3LYP-D3/def2-TZVP) and an implicit solvation model. While at a lower level of theory than the DHDF-D3 method used for the surface scans, this represents currently the highest-level QM calculation for a large Z-DNA structure done to date. The global conformation of

**Figure 6. Visualization of a full CpG step in a Z-DNA context illustrating the compression effect induced by the base pairing.** a) View of a fragment of the 3P4J Z-DNA X-ray structure (for a view of the full-duplex, see **Figure S1**). The 3.42 Å rise between C=G pairs, calculated between the two cytosine rings, corresponds roughly to the sum of the intranucleotide 0.53/0.63 Å O4'...nucleobase plane distance (grey zone) added to the 2.88/2.85 Å  $lp \dots \pi$  contact distance. b) Interaction energy curves for the DME...G system (see **Figure 4**) highlighting the projected compression zone (grey) and compression direction (dashed red arrow). Note that the interaction energies remain attractive in the 2.60-2.97 Å range.



Z-DNA is well retained as shown by the reasonably low (0.17 Å) DFT-D3 versus X-ray structure heavy root mean square deviation value (Figure S4). The  $2.83 \pm 0.03$  Å average lp... $\pi$  distance in the DFT-D3 structure meets the  $2.83 \pm 0.04$  Å X-ray average distance (Table S1). Even though a direct comparison between a crystal structure and an optimized model is problematic due to missing unit cell periodicity and the application of a continuum solvation model, the DFT-D3 structure is close to the X-ray structure. Therefore, it is necessary for an accurate geometrical description of the lp... $\pi$  contact to take into account the full Z-DNA structural context since the DME...G and *d*-ribose...G models are insufficient to capture the totality of the contributions leading to short contacts. Here, short lp... $\pi$  contacts originate from compression effects due to stereochemical constraints imposed by the specific left-handed Z-DNA backbone and base pair arrangement (Figure 6). In this lp... $\pi$  context, compression is facilitated by the shallow interaction energy profile. Therefore, pushing the O4' oxygen towards the guanine pyrimidine ring imposes a limited and tolerable local energy penalty of 0.1–0.2 kcal.mol<sup>-1</sup> over a 0.1 Å displacement for either model systems (DHDF-D3). The  $\approx 3.4$  Å interstrand base pair stacking distance calculated between the two central cytosines corresponds to a standard nucleic acid stacking distance and has to be preserved since it is known that deviations from an ideal vertical separation distance cause structural instabilities due to large gradients (forces).<sup>[19,28]</sup> Assuming such a stacking value, the only imaginable geometrical arrangement is to place the ribose O4' atom over, or slightly aside, the pyrimidine ring vdW depression.

## Discussion

We report on high level QM computations performed on the prototypical lp... $\pi$  system as it occurs in Z-DNA CpG steps. Our calculations reveal an intricate combination of interatomic interactions and rarely addressed fine conformational adjustments (summarized in Figure 7) that are at the origin of one of the shortest possible distance of approach of an oxygen atom to an aromatic ring. Our data may help better understand similar short lp... $\pi$  contacts in nucleic acids<sup>[6g]</sup> and other chemical systems. Relevant examples include lp... $\pi$  contacts in UNCG tetraloops that embed a CpG Z-step as well as the recurrent stacking of phosphate groups over nucleobases that occur in GNRA tetraloops or in the U-turn of tRNA anticodon loops, a motif essential for the decoding of the genetic information.<sup>[12,29]</sup>

Our data stress that unusually short lp... $\pi$  distances involving aromatic systems can hardly be understood without mentioning the vdW depression present at the center of these rings.<sup>[27]</sup> The position of the oxygen atom over this depression is reminiscent of the crystal packing principle of 'bumps and hollows' proposed by B.P. Orelkin in 1929/30, stating that bumps in one irregularly shaped molecular fragment fit into complementary hollows so that the molecules are closely packed with minimum voids between them.<sup>[16a,30]</sup> The fit of the O4' atom into such a hollow space is observed for the CpG steps in a full Z-DNA context and in *d*-ribose...G but is most prominently observed in DME...G where, in the absence of secondary interactions, the O4' atom sits exactly over the pyrimidine ring depression (Figure 5).

Yet, a full understanding of the short lp... $\pi$  distance in Z-steps also needs to integrate compression effects induced by the Z-DNA structural context. Since the lp... $\pi$  interactions are only weakly attractive when compared to the slightly stronger base stacking or to the much stronger base pairing interactions in nucleic acids, they are representative of 'soft' interactions displaying a high compressibility. This led us to revisit the introductory statements regarding short interatomic distances being indicators of attractive NCIs. In our opinion, the stereochemical effects described here, fully support the Dunitz and Gavezzotti (D&G) remarks that the latter assumptions may not be systematically valid. For instance, it is often wrongly assumed that the shortest hydrogen bonds are the strongest while it is established that the strongest are those closest to the optimal interaction geometry.<sup>[15a,15b]</sup> In that respect, D&G stated that "short contacts, often taken as evidence of unusual stabilization, may in fact be the result of compression from other parts of the crystal potential".<sup>[16c]</sup> D&G also discussed the occurrence of Coulombic compression effects in the context of X-ray structures of croconate salts where interlayer distances of aromatic rings as short as 3.1–3.4 Å result from cation-cation and anion-anion interaction energies being essentially compensated by cation-anion interaction energies.<sup>[17]</sup> Thus, although it is clearly impossible to dissect all interactions intervening in the stabilization of Z-DNA steps, the present lp... $\pi$  system illustrates that equilibrium within a complex molecular assembly results from a balance of all intermolecular forces rather than from all atom-pairs (or fragments) being at their respective minimum interaction energy distance.

This study stresses that the short lp... $\pi$  distances in Z-DNA steps and elsewhere are induced by the molecular context and do not originate from the weak lp... $\pi$  interaction. As such, it emphasizes a key aspect of NCI "ecosystems", i.e. that strong

**Figure 7.** Schematics showing the effects leading to short lp- $\pi$  contacts in CpG Z-steps. a) Estimated base stacking distance in B-DNA structures.<sup>[3]</sup> b) Oxygen...guanine distance according to the DME...G closest-contact surface. The oxygen atom takes advantage of the vdW depression of the pyrimidine ring (Figure 5). c) Oxygen...guanine distance according to the *d*-ribose...G closest contact surface (Figure 2). Stronger interaction energies involving secondary interactions lead to a shortening of the lp... $\pi$  contact distance. d) Oxygen...guanine distance in the X-ray CpG structure. Among others, base pair constraints induce vertical compression that leads to the observed shortening of the lp... $\pi$  contact distance (Figures 1/6).

and weak interactions cooperate to form the equilibrium structures we observe.<sup>[31]</sup>

From a methodological point of view, this study suggests that QM calculations on model systems may not always be able to accurately determine non-covalent interaction distances if these interactions, within the full structural context, are affected by compression or other structural effects.<sup>[32]</sup> However, it underlines, through a thorough comparison between high-quality calculated and experimental data, the usefulness of QM calculations in determining where such effects may be active.

It remains that one of the main surprises of this study concerns the understanding that the standard carbon vdW radii (1.70 Å)<sup>[2,13]</sup> that is widely used for assessing that  $lp \dots \pi$  interactions operate below the sum of the vdW radii of the intervening atoms<sup>[7]</sup> are significantly overestimated for  $sp^2$  atoms belonging to the highly polar guanidine system. Similar interrogations have been raised in other instances,<sup>[15c,15d]</sup> and it is sobering to recall that Bondi, in its 1964 original publication,<sup>[2]</sup> stated that his vdW radii may not be suitable for the calculation of contact distances in crystals. More precisely, his carbon vdW radius (1.70 Å) was designed for  $sp^3$  and not for  $sp^2$  atoms like those encountered in the Z-DNA guanidine fragment and other abundant chemical fragments such as, to cite a few, urea, guanidinium ions and carboxylate groups (see also the Alvarez<sup>[13]</sup> 1.77 Å  $sp^3$  carbon vdW radius). Thus, specific environmental adjustments of carbon vdW radii are needed to clear crystallographic surveys of short distances of approach misinterpretations.<sup>[25,33]</sup> They may also help better interpret the short contact distances of a nucleophile to a polar carbon exemplified by the Bürgi-Dunitz nucleophilic reaction path that is likewise characterized by short contact distances of a nucleophile to an  $sp^2$  carbon atom connected to polar groups.<sup>[34]</sup>

To conclude, this study challenges the whole  $lp \dots \pi$  interaction concept that is rooted in an inappropriate value of the carbon vdW radius that led to the assumption that these interactions belong to the sub-vdW category associated with orbital overlap characteristics.<sup>[35]</sup> In the light of our analysis of prototypical Z-steps, short  $lp \dots \pi$  interactions should be regarded as weak “secondary” interactions induced by locally strong structural constraints. Hence, we note that the  $lp \dots \pi$  interaction term does not accurately describe the forces that drive the formation of the ribose...guanine contact in Z-DNA and call for an overall reexamination of the  $lp \dots \pi$  concept as others urged rethinking the widely used ‘ $\pi \dots \text{stacking}$ ’ or ‘ $\pi \dots \pi$ ’ interaction terms they consider as misnomers.<sup>[36]</sup>

## Computational Methods

**Lone-pair- $\pi$  model preparation.** A deoxyribose...guanine fragment was extracted from the 0.55 Å ultra-high-resolution d(CpGpCpGpCpG)<sub>2</sub> Z-DNA structure (C3pG4 step; PDBid: 3P4J; see [Figures 1/S1](#)).<sup>[4]</sup> The ribose ring was transformed into dimethylether (DME) or appropriately saturated (*d*-ribose).

**Surface scan procedure.** Twenty-nine guanine nucleobase plane points (cartesian x-y coordinate plane, black dots in [Figure 2](#)), including the geometrical-centroid of the pyrimidine and imidazole rings, were chosen and 20 vertical (z-axis) displacements at 0.05 and 0.1 Å intervals between 2.5 and 4.2 Å were imposed to the oxygen (O4') atom of the DME and *d*-ribose probes (smaller steps were chosen around the expected

minimum; see [Supporting Information](#)). This resulted in a total of 29 x 20 = 580 single-point calculations per scan. The QM calculations were performed after optimization of the X-ray derived DME...G and *d*-ribose...G models by using geometrically and spatially frozen fragments to preserve the orientation of the oxygen atom lone pairs (hydrogen atoms were frozen; nucleobase planarity was enforced; see [Supporting Information](#)).

**Quantum mechanical (QM) calculations.** For the single-point energy calculations of the surface scan, we used the dispersion-corrected double-hybrid density functional approximation DSD-BYLP-D3<sup>[37]</sup> with the minimally-augmented quadruple- $\zeta$  basis set ma-def2-QZVPP<sup>[38]</sup> and BJ-damped DFT-D3 method<sup>[39]</sup> abbreviated as DHDF-D3. We benchmarked DHDF-D3 against CCSD(T)/CBS data, the current QM “gold-standard”,<sup>[40]</sup> and found only minimal deviations between both models that supported the use of DHDF-D3 for providing high accuracy results ([Supporting Information](#) and [Figures S3/S5](#)).

For the interaction energy decomposition, the symmetry-adapted perturbation theory (SAPT) was used at the SAPT2+3/cc-pVTZ level.<sup>[41]</sup> The SAPT2+3 approach includes important third-order corrections to the induction and dispersion terms that are often neglected.<sup>[41b]</sup> Note that the SAPT2+3 interaction energy values (abbreviated as SAPT) are slightly less accurate than the DHDF-D3 values. Charge-transfer effects (CT) were calculated following Stone and Misquitta<sup>[42]</sup>. Natural bond orbital (NBO)<sup>[24]</sup> computations used a B3LYP/def2-QZVP wave function.

The large-scale Z-DNA hexamer duplex calculations were performed at the less “costly” COSMO-B3LYP-D3/def2-TZVP (DFT-D3) level of theory. The COSMO model constructs a charged cavity around the molecule that screens the excess anionic charge of the Z-DNA. Thus, explicit use of counter ions is not required. Additional computational details are given in the [Supporting Information](#).

**Lone-pair... $\pi$  contact distances and stacking distances.** We underline that our calculated  $lp \dots \pi$  distances correspond to oxygen...nucleobase plane and not to oxygen...pyrimidine ring centroid distances as often seen in the literature. To quantify nucleobase stacking distances, the average of all pyrimidine ring-atom distances to the opposing nucleobase plane was calculated, i.e. each stacking value is an average over a total of 12 point-to-plane distances ([Table S1](#) and [Supporting Information](#)).

## Acknowledgements

The authors wish to thank Dr. Quentin Vicens, Prof. Eric Westhof and Prof. Georges Wipff for careful reading of the manuscript and for constructive discussions. We especially thank Dr. Maura Malinska for having provided the density maps issued from a TAAM refinement of the 3P4J structure.<sup>[1]</sup> HK and JS were supported by the project SYMBIT reg. number: CZ.02.1.01/0.0/0.0/15\_003/0000477 financed by the ERDF.

**Keywords:**  $lp \dots \pi$  interaction • molecular recognition • non-covalent interaction • van der Waals radii • Z-DNA

- [1] M. Malinska, Z. Dauter, Transferable aspherical atom model refinement of protein and DNA structures against ultrahigh-resolution X-ray data. *Acta Cryst.* **2016**, D72, 770-779.
- [2] A. Bondi, van der Waals volumes and radii. *J. Chem. Phys.* **1964**, 68, 441-451.
- [3] R. Chandrasekaran, S. Arnott, The structure of B-DNA in oriented fibers. *J. Biomol. Struct. Dyn.* **1996**, 13, 1015-1027.
- [4] K. Brzezinski, A. Brzuszkiewicz, M. Dauter, M. Kubicki, M. Jaskolski, Z. Dauter, High regularity of Z-DNA revealed by ultra high-resolution crystal structure at 0.55 Å. *Nucleic Acids Res.* **2011**, 39, 6238-6248.
- [5] a) L. D'Ascenzo, F. Leonarski, Q. Vicens, P. Auffinger, 'Z-DNA like' fragments in RNA: a recurring structural motif with implications for folding, RNA/protein recognition and immune response. *Nucleic Acids Res.* **2016**, 44, 5944-5956; b) A. Bauza, T. J. Mooibroek, A. Frontera, The bright future of unconventional sigma/pi-hole interactions. *Chemphyschem* **2015**, 16, 2496-2517.
- [6] a) M. Egli, R. V. Gessner, Stereoelectronic effects of deoxyribose O4' on DNA conformation. *Proc. Natl. Acad. Sci. USA* **1995**, 92, 180-184; b) E. A. Meyer, R. K. Castellano, F. Diederich, Interactions with aromatic rings in chemical and biological recognition. *Angew. Chem. Int. Ed. Engl.* **2003**, 42, 1210-1250; c) M. Egli, S. Sarkhel, Lone pair-aromatic interactions: to stabilize or not to stabilize. *Acc. Chem. Res.* **2007**, 40, 197-205; d) L. M. Salonen, M. Ellermann, F. Diederich, Aromatic rings in chemical and biological recognition: energetics and structures. *Angew. Chem. Int. Ed. Engl.* **2011**, 50, 4808-4842; e) N. Mohan, C. H. Suresh, A. Kumar, S. R. Gadre, Molecular electrostatics for probing lone pair- $\pi$  interactions. *Phys. Chem. Chem. Phys.* **2013**, 15, 18401-18409; f) S. K. Singh, A. Das, The  $n \rightarrow \pi^*$  interaction: a rapidly emerging non-covalent interaction. *Phys. Chem. Chem. Phys.* **2015**, 17, 9596-9612; g) J. Novotny, S. Bazzi, R. Marek, J. Kozelka, Lone pair- $\pi$  interactions: analysis of the physical origin and biological implications. *Phys. Chem. Chem. Phys.* **2016**, 18, 19472-19481; h) M. Chawla, E. Chermak, Q. Zhang, J. M. Bujnicki, R. Oliva, L. Cavallo, Occurrence and stability of lone pair- $\pi$  stacking interactions between ribose and nucleobases in functional RNAs. *Nucleic Acids Res.* **2017**, 45, 11019-11032; i) J. Kozelka, Lone pair- $\pi$  interactions in biological systems: occurrence, function, and physical origin. *Eur. Biophys. J.* **2017**, 46, 729-737.
- [7] T. J. Mooibroek, P. Gamez, J. Reedijk, Lone pair- $\pi$  interactions: a new supramolecular bond? *CrystEngComm* **2008**, 10, 1501-1515.
- [8] a) A. J. Neel, M. J. Hilton, M. S. Sigman, F. D. Toste, Exploiting non-covalent  $\pi$  interactions for catalyst design. *Nature* **2017**, 543, 637-646; b) Y. Zhao, Y. Cotelte, L. Liu, J. Lopez-Andarias, A. B. Bornhof, M. Akamatsu, N. Sakai, S. Matile, The emergence of anion- $\pi$  catalysis. *Acc. Chem. Res.* **2018**, 51, 2255-2263.
- [9] J. Sponer, H. A. Gabb, J. Leszczynski, P. Hobza, Base-base and deoxyribose-base stacking interactions in B-DNA and Z-DNA: a quantum-chemical study. *Biophys. J.* **1997**, 73, 76-87.
- [10] G. Knizia, Intrinsic atomic orbitals: an unbiased bridge between quantum theory and chemical concepts. *J. Chem. Theory Comput.* **2013**, 9, 4834-4843.
- [11] A. H. J. Wang, G. J. Quigley, F. J. Kolpak, G. Vandermaarel, J. H. Vanboom, A. Rich, Left-handed double helical DNA - Variations in the backbone conformation. *Science* **1981**, 211, 171-176.
- [12] L. D'Ascenzo, F. Leonarski, Q. Vicens, P. Auffinger, Revisiting GNRA and UNGC folds: U-turns versus Z-turns in RNA hairpin loops. *RNA* **2017**, 23, 259-269.
- [13] S. Alvarez, A cartography of the van der Waals territories. *Dalton Trans.* **2013**, 42, 8617-8636.
- [14] M. Savastano, C. Garcia-Gallarin, M. D. Lopez de la Torre, C. Bazzicalupi, A. Bianchi, M. Melguizo, Anion- $\pi$  and lone pair- $\pi$  interactions with s-tetrazine-based ligands. *Coord. Chem. Rev.* **2019**, 397, 112-137.
- [15] a) T. Steiner, The hydrogen bond in the solid state. *Angew. Chem. Int. Ed.* **2002**, 41, 48-76; b) I. Dance, Distance criteria for crystal packing analysis of supramolecular motifs. *New J. Chem.* **2003**, 27, 22-27; c) G. P. Schiemenz, The sum of van der Waals radii - A pitfall in the search for bonding. *Z. Naturforsch.* **2007**, 62b, 235-243; d) S. Z. Hu, Z. H. Zhou, Z. X. Xie, B. E. Robertson, A comparative study of crystallographic van der Waals radii. *Z. Kristallogr.* **2014**, 229, 517-523.
- [16] a) J. D. Dunitz, A. Gavezzotti, Attractions and repulsions in molecular crystals: what can be learned from the crystal structures of condensed ring aromatic hydrocarbons. *Acc. Chem. Res.* **1999**, 32, 677-684; b) J. D. Dunitz, A. Gavezzotti, Molecular recognition in organic crystals: directed intermolecular bonds or nonlocalized bonding? *Angew. Chem. Int. Ed. Engl.* **2005**, 44, 1766-1787; c) J. D. Dunitz, A. Gavezzotti, How molecules stick together in organic crystals: weak intermolecular interactions. *Chem. Soc. Rev.* **2009**, 38, 2622-2633; d) A. Gavezzotti, The lines-of-force landscape of interactions between molecules in crystals: cohesive versus tolerant and 'collateral damage' contact. *Acta Cryst.* **2010**, B66, 396-406; e) J. D. Dunitz, Intermolecular atom-atom bonds in crystals? *IUCrJ* **2015**, 2, 157-158.
- [17] J. D. Dunitz, A. Gavezzotti, S. Rizzato, "Coulombic compression", a pervasive force in ionic solids. A study of anion stacking in croconate salts. *Cryst. Growth Des.* **2014**, 14, 357-366.
- [18] F. Leonarski, L. D'Ascenzo, P. Auffinger,  $Mg^{2+}$  ions: do they bind to nucleobase nitrogens? *Nucleic Acids Res.* **2017**, 45, 987-1004.
- [19] H. Kruse, P. Banas, J. Sponer, Investigations of stacked DNA base-pair steps: highly accurate stacking interaction energies, energy decomposition, and many-body stacking effects. *J. Chem. Theory Comput.* **2019**, 15, 95-115.
- [20] E. G. Hohenstein, C. D. Sherill, Wavefunction methods for noncovalent interactions. *WIREs Comput. Mol. Sci.* **2012**, 2, 304-326.
- [21] G. Gryn'ova, C. Corminboeuf, Steric "attraction": not by dispersion alone. *Beilstein J. Org. Chem.* **2018**, 14, 1482-1490.
- [22] T. Bitter, K. Ruedenberg, W. H. Schwarz, Toward a physical understanding of electron-sharing two-center bonds. I. General aspects. *J. Comput. Chem.* **2007**, 28, 411-422.
- [23] P. Jurecka, J. Sponer, J. Cerny, P. Hobza, Benchmark database of accurate (MP2 and CCSD(T) complete basis set limit) interaction energies of small model complexes, DNA base pairs, and amino acid pairs. *Phys. Chem. Chem. Phys.* **2006**, 8, 1985-1993.
- [24] E. D. Glendenning, C. R. Landis, F. Weinhold, Natural bond orbital methods. *WIREs Comput. Mol. Sci.* **2012**, 2, 1-42.
- [25] R. Taylor, P. A. Wood, A million crystal structures: The whole is greater than the sum of its parts. *Chem. Rev.* **2019**.
- [26] P. Politzer, P. Jin, J. S. Murray, Atomic polarizability, volume and ionization energy. *J. Chem. Phys.* **2002**, 117, 8197-8202.
- [27] B. W. Gung, Y. Zou, Z. G. Xu, J. C. Amicangelo, D. G. Irwin, S. Q. Ma, H. C. Zhou, Quantitative study of interactions between oxygen lone pair and aromatic rings: Substituent effect and the importance of closeness of contact. *J. Org. Chem.* **2008**, 73, 689-693.
- [28] T. M. Parker, E. G. Hohenstein, R. M. Parrish, N. V. Hud, C. D. Sherill, Quantum-mechanical analysis of the energetic contributions to pi stacking in nucleic acids versus rise, twist, and slide. *J. Am. Chem. Soc.* **2013**, 135, 1306-1316.
- [29] P. Auffinger, E. Westhof, An extended structural signature for the tRNA anticodon loop. *RNA* **2001**, 7, 334-341.
- [30] L. Glasser, in *Fundamental Principles of Molecular Modeling*. (Eds.: W. Gans, A. Amann, J. C. A. Boeyens), Springer, New-York, **1996**.
- [31] a) P. Csermely, Strong links are important, but weak links stabilize them. *Trends Biochem. Sci.* **2004**, 29, 331-334; b) A. S. Batsanov, Weak interactions in crystals: old concepts, new developments. *Acta Cryst.* **2018**, E74, 570-574.
- [32] O. Khakhsheer, S. E. Wheeler, K. N. Houk, E. T. Kool, Measurement and theory of hydrogen bonding contribution to isosteric DNA base pairs. *J. Am. Chem. Soc.* **2012**, 134, 3154-3163.
- [33] H. Kruse, J. Sponer, P. Auffinger, Comment on "Evaluating Unexpectedly Short Non-covalent Distances in X-ray Crystal Structures of Proteins with Electronic Structure Analysis". *J. Chem. Inf. Model.* **2019**, 59, 3605-3608.
- [34] a) H. B. Bürgi, J. D. Dunitz, E. Shefter, Geometrical reaction coordinates. II. Nucleophilic addition to a carbonyl group. *J. Am. Chem. Soc.* **1973**, 95, 5065-5067; b) H. B. Bürgi, J. D. Dunitz, J. M. Lehn, G. Wipff, Stereochemistry of reaction paths at carbonyl centres. *Tetrahedron* **1974**, 30, 1563-1572.
- [35] B. Worley, G. Richard, G. S. Harbison, R. Powers,  $^{13}C$  NMR reveals no evidence of  $n-\pi^*$  interactions in proteins. *PLoS One* **2012**, 7, e42075.
- [36] a) S. Grimme, Do special noncovalent  $\pi-\pi$  stacking interactions really exist? *Angew. Chem. Int. Ed. Engl.* **2008**, 47, 3430-3434; b) C. R. Martinez, B. L. Iverson, Rethinking the term "pi-stacking". *Chem. Sci.* **2012**, 3, 2191-2201.
- [37] S. Kozuch, D. Gruzman, J. M. L. Martin, A general purpose double hybrid density functional including spin component scaling and dispersion correction. *J. Phys. Chem.* **2010**, 114, 20801-20808.
- [38] J. Zheng, X. Xu, D. G. Truhlar, Minimally augmented Karlsruhe basis sets. *Theor. Chem. Acc.* **2011**, 128, 295-305.
- [39] S. Grimme, S. Ehrlich, L. Goerigk, Effect of the damping function in dispersion corrected density functional theory. *J. Comput. Chem.* **2011**, 32, 1456-1465.
- [40] L. A. Burns, M. S. Marshall, C. D. Sherill, Appointing silver and bronze standards for noncovalent interactions: a comparison of spin-component-scaled (SCS), explicitly correlated (F12), and specialized wavefunction approaches. *J. Chem. Phys.* **2014**, 141, 234111.
- [41] a) E. G. Hohenstein, C. D. Sherill, Efficient evaluation of triple excitations in symmetry-adapted perturbation theory via second-order Møller-Plesset perturbation theory natural orbitals. *J. Chem. Phys.* **2010**, 133, 104107; b) T. M. Parker, L. A. Burns, R. M. Parrish, A. G. Ryno, C. D. Sherill, Levels of symmetry adapted perturbation theory (SAPT). I. Efficiency and performance for interaction energies. *J. Chem. Phys.* **2014**, 140, 094106.
- [42] A. J. Stone, A. J. Misquitta, Charge-transfer in symmetry-adapted perturbation theory. *Chem. Phys. Lett.* **2009**, 473, 201-205.



WILEY-VCH